

Ähnlichkeitssuche in den Digital Humanities: Semi-automatische Identifikation von Kostümmustern

Johanna Barzen¹, Michael Falkenthal¹, Frank Hentschel², Frank Leymann¹, Tino Strehl³

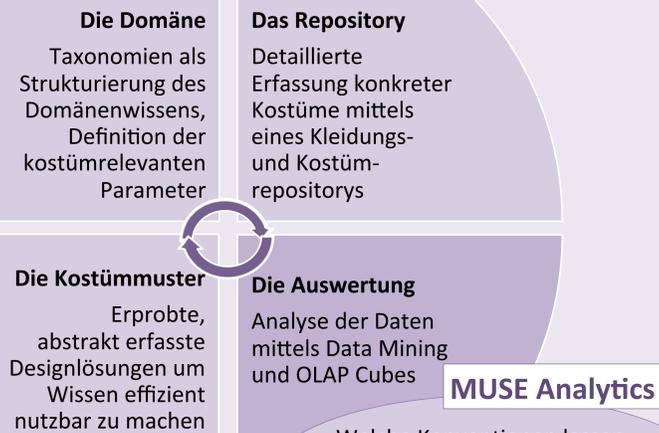
1. Institut für Architektur von Anwendungssystemen, Universität Stuttgart

2. Musikwissenschaftliches Institut, Universität zu Köln

3. Herman Hollerith Zentrum, Hochschule Reutlingen

MUSE – Muster Suchen und Erkennen

Die MUSE Methode



Die Ausgangslage:

MUSE ist eine Werkzeugumgebung, um Kostümmuster zu identifizieren. Ein Kostümmuster ist eine bewährte Lösung eines wiederkehrenden Kostümdesignproblems. Beispielsweise um einen adäquaten textilen Ausdruck eines bestimmten Charakters zu finden. Die *MUSE Methode* beinhaltet die detaillierte Erfassung von Daten, wie auch deren Analyse, um Hinweise auf Kostümmuster zu identifizieren und diese zu Mustern zu abstrahieren. Um Hinweise auf Muster zu ermöglichen besteht *MUSE Analytics* aus sich abwechselnden Data Mining- und Validierungsphasen (OLAP Cube). Der hier vorgestellte Ansatz veranschaulicht wie die taxonomische Strukturierung der Daten genutzt werden kann, um mittels Ähnlichkeitsanalyse Cluster zu identifizieren, welche durch verschiedene Visualisierungsverfahren Hinweise auf mögliche Muster geben können:

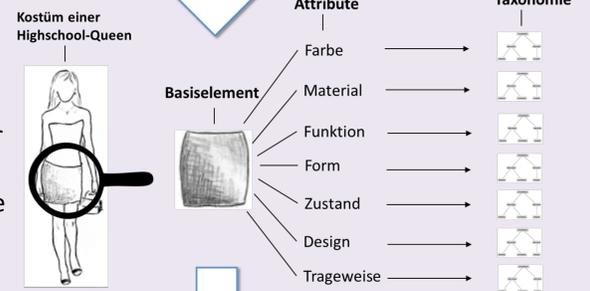
Die Daten

Konkrete Kostüme aus Filmen (Filmkorpus) werden detailliert mit Hilfe von Eingabemasken und vordefinierten kostümrelevanten Parametern erfasst.



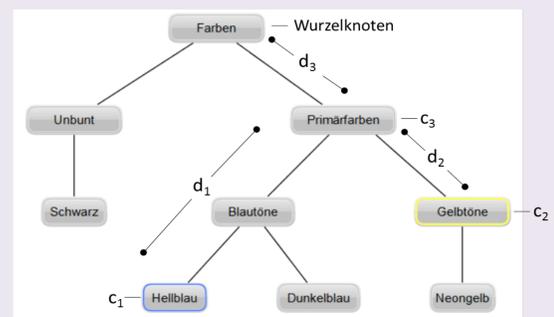
Die Datenstruktur

Die kostümrelevanten Parameter sind taxonomisch strukturiert: Jedes Kostüm und seine Basiselemente werden durch Attribute der Taxonomien (wie beispielsweise Farben, Materialien, Designs etc.) detailliert und präzise beschrieben.



Das Grundprinzip der Ähnlichkeitsanalyse

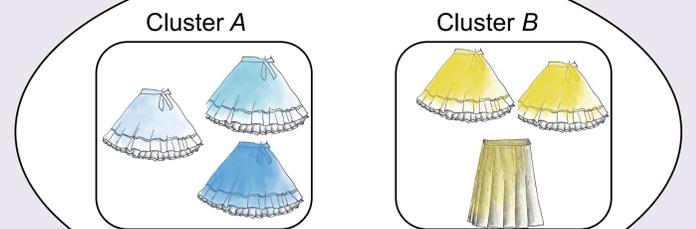
Auf Basis der strukturellen Eigenschaften der Taxonomien können die semantischen Ähnlichkeit zwischen den Attributen bestimmt werden. Die Ähnlichkeitsbestimmung erfolgt durch das Prinzip der Kantenzählung nach der Wu und Palmer Metrik.



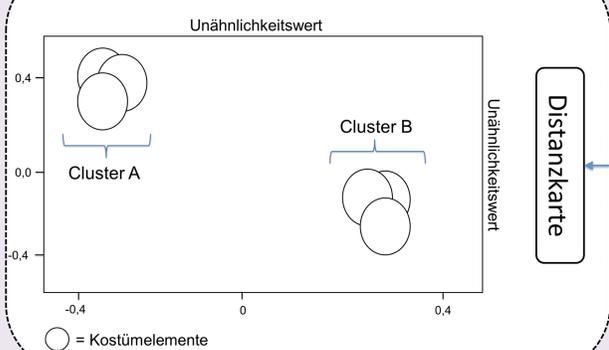
$$\text{ConSim}(\text{Hellblau}, \text{Gelbtöne}) = \frac{2 * d_3}{d_1 + d_2 + (2 * d_3)}$$

Das Clustering

Gruppierung der Basiselemente nach den berechneten semantischen Ähnlichkeitswerten auf Basis des hierarchischen Clusterverfahrens.



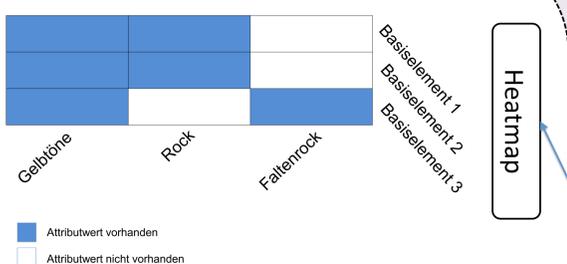
Elemente des Datenraums



Visualisierung der Elemente

Distanzkarte

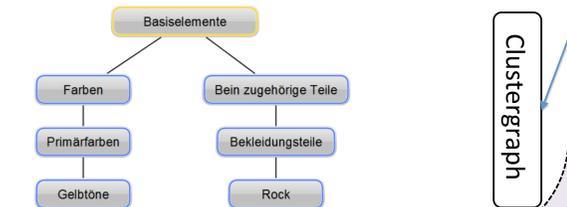
Werte des Cluster B



Visualisierung der Werte

Heatmap

Werte(-struktur) des Cluster B



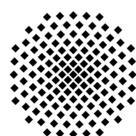
Die Visualisierung

Die Visualisierung dient der Identifikation von potentiellen Kostümmustern. Um die Suche nach Hinweisen zu unterstützen, werden unterschiedliche Visualisierungsverfahren genutzt: Hier z.B. die graphenbasierte und mapbasierte Visualisierung für die Darstellung der Werte der Basiselemente eines Clusters oder die Darstellung der Basiselemente durch einen Positionierungsalgorithmus auf Basis der Ähnlichkeitswerte.

Weitere Informationen

Ausgewählte Publikationen

- Schumm, D.; Barzen, J.; Leymann, F.; Ellrich, L.: A Pattern Language for Costumes in Films. In: Proceedings of the 17th European Conference on Pattern Languages of Programs (EuroPLoP 2012).
- Barzen, J.; Leymann, F.: Costume Languages As Pattern Languages. In: Proceedings of PURPLSOC (Pursuit of Pattern Languages for Societal Change). The Workshop 2014.
- Falkenthal, M.; Barzen, J.; Breitenbücher, U.; Fehling, C.; Leymann, F.: Efficient Pattern Application: Validating the Concept of Solution Implementations in Different Domains. In: International Journal On Advances in Software Vol. 7 (3&4), Xpert Publishing Services, 2014.



Universität Stuttgart



Kontakt

Johanna Barzen

Universität Stuttgart

Institut für Architektur von Anwendungssystemen

Universitätsstrasse 38

70569 Stuttgart

Phone: +49 711 685-88 487

Email: Barzen@iaas.uni-stuttgart.de